光电技术及应用

DOI：10.16818/j.issn1001－5868.20250226002

新型电力系统中数字孪生辅助的网络切片资源分配

唐奕明*，刘崇茹

（华北电力大学 新能源电力系统国家重点实验室，北京 102206）

摘　要： 针对新型电力系统中服务功能链（SFC）部署面临的资源需求动态变化与低时延需求，文章提出一种数字孪生辅助的虚拟网络功能（VNF）映射与关联优化算法。首先，通过构建VNF数字孪生体实时采集资源需求信息，建立以最小化SFC端到端时延和VNF平均数字孪生同步时延为目标的马尔科夫决策过程（MDP）模型，并考虑计算、存储及带宽资源约束。其次，将原问题分解为VNF映射与孪生服务节点关联两个阶段，分别提出基于分布式近端策略优化（DPPO）的VNF映射算法和基于优势演员-评论家（A2C）的孪生服务节点关联算法，以应对大规模SFC场景下的动作空间爆炸与策略波动问题。仿真结果表明，所提算法能够有效降低SFC端到端时延及VNF数字孪生同步时延，在多业务动态环境下具备良好的收敛性与资源适配能力。

关键词： 新型电力系统；虚拟网络功能；服务功能连端到端时延；VNF平均数字孪生同步时延

中图分类号： TN915；TM734　**文章编号：** 1001－5868（2026）02－0407－10

Software-Defined Network Slicing Test Bench and Validation for White-Box Switches

TANG Yiming, LIU Chongru

（**State Key Laboratory of Alternate Electrical Power System with Renewable Energy Source, North China Electric Power University, Beijing 102206, CHN**）

Abstract: The deployment of service function chains (SFCs) in modern power systems faces considerable challenges owing to dynamically changing resource demands and stringent low-latency requirements. To address these challenges, this paper presents a digital twin-assisted optimization algorithm for virtual network function (VNF) mapping and association. Initially, VNF digital twins were constructed to capture real-time resource demand information, thus enabling the formulation of a Markov decision process (MDP) model. The optimization objective was to minimize both the end-to-end delay of the SFCs and the average synchronization delay of the VNF digital twins while satisfying the constraints on computing, storage, and bandwidth resources. Given the sequential nature of VNF mapping and twin service node association, the original problem was decomposed into two stages. A distributed proximal policy optimization (DPPO)-based algorithm was proposed for the VNF mapping stage, and an advantage actor-critic (A2C)-based algorithm was developed for the twin-service node association stage. These algorithms effectively mitigate the issues of action-space explosion and considerable policy fluctuations encountered by heuristic and basic reinforcement learning methods as the number of SFCs increases. The simulation results demonstrate that the proposed DPPO-A2C algorithm can substantially reduce the SFC end-to-end and VNF digital twin synchronization delays by

收稿日期：2025－02－26.

*通信作者：唐奕明

designing appropriate mapping and association schemes, exhibiting superior convergence and resource adaptation performance in dynamic multiservice scenarios.

Key words: modern power system; virtual network function; service function chain end-to-end delay; VNF average digital twin synchronization delay

0 引言

随着智能电网和配电网系统的深入建设,新型电力系统亟需高带宽、高可靠、低时延的移动通信网络[1]。网络切片作为第五代移动通信技术(5G)的关键技术,借助网络功能虚拟化(NFV)和软件定义网络(SDN)技术使得网络能够满足不同配电网服务的差异化需求[2]。

目前,随着电力物联网的快速发展,新型电力系统切片业务急剧增加,同时也对网络切片的资源优化能力提出了更高要求。为了满足不同服务的动态资源需求特性与时延需求,需要对电力网络中的资源进行合理分配,并对新型电力系统中的网络切片资源分配方案进行深入研究。文献[3]首先结合5G切片技术,构建了配电自动化、差动保护和精准负控等电力专属网络应用场景,然后分析了电力5G切片给电网带来的经济效益,最后从技术应用和经济效益两个层面论证电力5G切片赋能城市配电网的可行性。文献[4]研究了针对配电网多业务场景下的5G切片部署问题,构建了切片部署成本和时延模型,并提出了基于成本惩罚函数和时延惩罚函数的成本时延联合优化算法,该算法可以根据配电网业务终端的接入情况制定相应的切片部署策略。文献[5]研究了适用于配电网的5G网络切片技术,以提高通信的实时性,同时研究了基于边缘计算的相邻区域的自适应相关系数算法,以提高数据分析的稳定性。文献[6]提出遗传算法来解决网络切片的资源分配问题,以实现网络效益的最大化。

然而,通过当前的研究现状可以看出,大部分资源分配方案并未充分考虑新型电力系统的特殊场景,而只是将5G网络切片方案与新型电力系统进行简单的匹配,无法有效解决新型电力系统的网络切片资源分配问题。因此,本文针对新型电力系统中服务的动态资源需求变化特性,考虑将数字孪生技术引入到该系统中,通过VNF数字孪生体实时采集并更新服务的资源需求,并对VNF进行监管来保证VNF映射的可靠性。数字孪生起源于航天领域,近年来持续向汽车制造、工业控制、能源应用等多个垂直领域拓展[7-8],逐渐成为了助力企业实现数字化转型,促进数字经济发展的总抓手。文献[9]为数字孪生网络白皮书,它系统阐述了数字孪生网络的总体架构、关键技术和评估分级体系,明确了数字孪生网络是解决未来网络窘境,实现网络智能化的关键。通过数字孪生网络可以实现对物理网络的拓扑透视与流量全息;同时,设计具备人工智能能力的网络资源分配可以有效降低服务的端到端时延。此外,降低VNF到其数字孪生体的同步时延也至关重要,同步时延越小,新型电力系统就能更早获得资源需求设计VNF映射方案,防止VNF映射策略与服务不匹配造成的映射策略滞后。

因此,本文针对新型电力系统中服务资源需求的动态变化与低时延需求,提出一种数字孪生辅助的VNF映射算法,主要内容为:

1)针对新型电力系统的动态资源需求变化特性,设计了一种数字孪生辅助的VNF映射架构。

2)将随机优化问题建模为马尔科夫决策过程(MDP)模型,该模型以优化SFC端到端时延和VNF平均数字孪生同步时延为目标,同时受限于SFC端到端时延、通用服务器CPU、存储资源以及链路带宽资源等约束。

3)考虑到VNF映射和孪生服务节点具备先后顺序,将原问题分为VNF映射和孪生服务节点关联两个阶段进行求解,并提出了基于DPPO的VNF映射算法和基于A2C的孪生服务节点关联算法,其中基于DPPO的VNF映射算法能够解决DQN算法VNF映射幅度过大和PPO算法在SFC数量较多时动作搜索空间过大的问题。基于A2C算法能够衡量所选动作相较于所有动作平均值的好坏。

4)仿真结果表明,所提DPPO-A2C算法能够通过设计合理的VNF映射和关联方案来降低SFC的端到端时延和所有VNF的平均数字孪生同步时延。

1　新型电力系统中数字孪生辅助的网络切片资源分配模型

1.1　数字孪生辅助的网络切片资源分配模型

1.1.1　模型场景

针对新型电力系统中服务的动态变化特性和低延迟需求，本文提出了数字孪生辅助的网络切片资源分配模型，如图 1 所示，主要包括新型电力系统物理设施层、网络切片层与网络应用层。

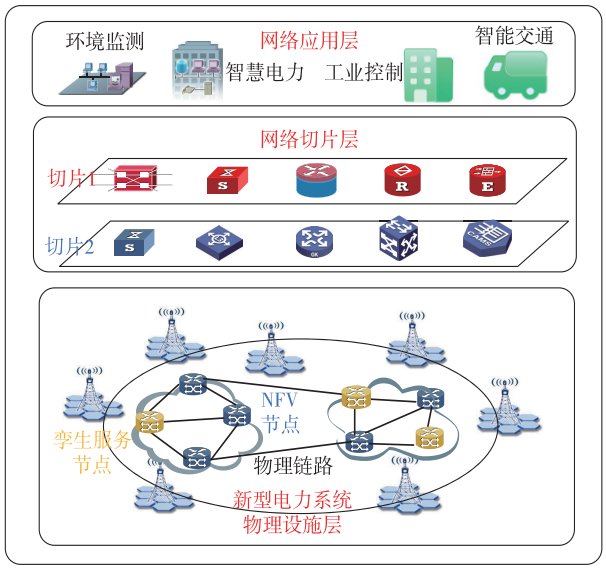


图 1　网络场景图

1.1.2　物理网络

将支持新型电力系统的物理网络建模为全连接无向图 $G_{\mathrm{P}}=(N_{\mathrm{P}},L_{\mathrm{P}})$，其中 N_{P} 为底层物理节点集，L_{P} 为物理链路集。物理节点既可用于处理 VNF，又可用于构建 VNF 数字孪生体，但一个节点只能单纯部署 VNF 实例或构建 VNF 数字孪生体。为此，本文将节点分为两类，即处理 VNF 实例的 NFV 功能节点集 $N_{\mathrm{P}}^{\mathrm{N}}$ 和构建 VNF 数字孪生体的节点集 $N_{\mathrm{P}}^{\mathrm{DT}}$，且 $N_{\mathrm{P}}^{\mathrm{N}}\cup N_{\mathrm{P}}^{\mathrm{DT}}=N_{\mathrm{P}}$。并设 C_n 和 M_n 分别表示 NFV 节点 $n\in N_{\mathrm{P}}^{\mathrm{N}}$ 的计算和存储资源，C_l^{DT} 和 M_l^{DT} 分别表示孪生服务节点 $l\in N_{\mathrm{P}}^{\mathrm{DT}}$ 的计算和存储资源，B_{mn} 表示物理链路 $mn\in L_{\mathrm{P}}$ 的带宽资源。

1.1.3　SFC 请求

设网络切片层中有一组 SFC 请求集，定义为 F。第 i 条 SFC 可抽象为有向图 $G_{\mathrm{V}}^i=(N_{\mathrm{V}}^i,L_{\mathrm{V}}^i)$，其中 N_{V}^i 为 SFC_i 上的 VNF 集合，L_{V}^i 为 SFC_i 上的虚拟链路集合。一条 SFC 上可能包含多个 VNF，设 $N_{\mathrm{V}}^{i,j}$ 表示第 i 条 SFC 上的第 j 个 VNF。此外，在底层物理网络中部署 SFC 会消耗一定的计算、存储和带宽资源，设 $C_n^{i,j}$ 和 $M_n^{i,j}$ 分别表示 SFC_i 上第 j 个 VNF 映射到节点 n 上的计算和存储资源需求。$B_{mn}^{i,jk}$ 表示 SFC_i 上连接相邻 VNF_{jk} 的虚拟链路带宽资源需求。接下来定义二进制变量 $x_n^{i,j}$ 和 $y_{mn}^{i,jk}$ 分别表示 SFC 上 VNF 和虚拟链路映射到物理网络的情况。当 SFC_i 上的 VNF_j 映射到物理节点 n 上时，$x_n^{i,j}=1$，否则 $x_n^{i,j}=0$。同理，当 VNF_{jk} 的虚拟链路映射到物理链路 mn 上时，$y_{mn}^{i,jk}=1$，否则 $y_{mn}^{i,jk}=0$。

1.1.4　时延模型

在计算 SFC 端到端时延的时候，主要考虑处理时延、排队时延、传输时延和传播时延。首先，物理节点运行器上部署的 VNF 需要处理相应的数据包，而处理数据包花费的时间为处理时延。处理时延主要与 VNF 传输的数据包和物理节点分配给它的计算资源有关，设 $p_{i,j-1}(t)$ 表示 SFC_i 上到达 VNF_j 的数据包，κ 表示数据包的处理速率系数，那么物理节点处理 VNF_j 的处理时延 $d_{\mathrm{p}}^{i,j}(t)$ 可表示为：

$$d_{\mathrm{p}}^{i,j}(t)=\frac{p^{i,j-1}(t)}{x_n^{i,j}(t)\cdot C_n^{i,j}(t)\cdot\kappa} \tag{1}$$

其次，排队时延表示 SFC 上 VNF 数据包达到需要排队等候的时间。对 SFC_i 而言，设 $R^i(t)$ 表示其在 t 时隙的队列长度，数据包的到达过程为 $A^i(t)$，且服从参数为 λ_i 的泊松分布[10]。则队列 $R^i(t)$ 的更新过程可表示为：

$$R^i(t+1)=\max\left[R^i(t)-x_n^{i,1}(t)\cdot C^{i,1}(t)\cdot\kappa+A^i(t),0\right] \tag{2}$$

而 SFC 的排队时延 $d_{\mathrm{q}}^i(t)$ 可表示为：

$$d_{\mathrm{q}}^i(t)=\frac{R^i(t)}{\lambda^i(t)} \tag{3}$$

然后，传输时延是指传输完数据包所花费的时间，主要与传输的数据量和分配给虚拟链路的带宽资源有关。设 $d_{\mathrm{tran}}^{i,jk}(t)$ 指 SFC_i 上虚拟链路 VNF_{jk} 之间的传输时延，那么其可表示为：

$$d_{\mathrm{tran}}^{i,jk}(t)=\frac{p^{i,j-1}(t)}{y_{mn}^{i,jk}(t)\cdot B_{jk}^i(t)} \tag{4}$$

最后，传播时延为数据包在物理链路介质中传播的时间，其主要取决于物理链路的距离。设 SFC 上虚拟链路 VNF_{jk} 映射到物理链路 mn，且 mn 的传输距离为 d_{mn}，那么链路 mn 的传播时延 $d_{\mathrm{pro}}^{i,jk,mn}(t)$ 可表示为 $d_{\mathrm{pro}}^{i,jk,mn}(t)=y_{mn}^{i,jk}(t)\cdot d_{mn}/c$，其中 c 为光速。

综上所述，SFC 部署的端到端时延为处理时延、排队时延、传输时延与传播时延之和。因此

SFC_i的端到端时延$d^i(t)$可表示为：

$$d^i(t)=d_{\mathrm{q}}^i(t)+\sum_{j,k\in N_{\mathrm{V}}^i}(d_{\mathrm{p}}^{i,j}(t)+d_{\mathrm{trans}}^{i,jk}(t)+d_{\mathrm{pro}}^{i,jk,mn}(t)),\forall mn\in L_{\mathrm{V}}^i \tag{5}$$

1.1.5 数字孪生体及其关联

在本节中，设系统在孪生服务节点上构建VNF数字孪生体，一方面便于实时采集和监控VNF的运行信息，另一方面为SFC的全局数据信息同步提供有效支撑。首先，设VNF数字孪生体可被建模为$DT_{ij}=\{D,M,W\}$，其中D表示VNF数字孪生体的历史数据，M表示物理节点上中VNF的数字模型，W表示该数字孪生模型的模型参数。设$z_{s,l}^{i,j}(t)$表示t时隙被映射到NFV节点s上的第j个VNF是否关联到孪生服务节点l，$z_{s,l}^{i,j}(t)=1$表示VNF_j关联到孪生节点l，即在孪生服务节点l上构建VNF_j的数字孪生体，反之不构建。其次，设$c_l^{i,j}$和$m_l^{i,j}$分别表示t时隙在孪生服务节点l上构建VNF_j数字孪生体的计算和存储资源需求，$b_{mn}^{i,j}$表示t时隙VNF_j在关联到其孪生服务节点的关联路径上物理链路mn的带宽资源需求。

1.1.6 数字孪生同步时延模型

在新型电力系统的基础设施层，假设各孪生服务节点将VNF资源需求信息传输到SDN控制器的时延远小于VNF到其数字孪生体的同步时延，可忽略不计。那么在VNF部署到相应的物理节点后，为了使SDN控制器快速获得VNF的实时资源需求信息，需要降低VNF到其数字孪生体的数字孪生同步时延。同时考虑到从部署VNF的NFV节点到其孪生服务节点可能存在多条路径，为了降低问题的复杂性，对SFC_i上的VNF_j来说，设从NFV节点s到孪生服务节点l选择时延最短的链路进行关联，称作关联链路$L_{s,l}^{i,j}$。而数字孪生同步时延主要由关联链路传输时延、关联链路传播时延和数字孪生构建时延三部分组成。首先，对于关联链路传输时延$\tau_{mn,\mathrm{tran}}^{i,j}(t)$，其主要与$VNF_j$传输的数据量$D_{ij}(t)$和物理链路分配给其的带宽资源有关，可表示为：

$$\tau_{mn,\mathrm{tran}}^{i,j}(t)=\frac{D_{ij}(t)}{b_{mn}^{i,j}(t)} \tag{6}$$

式中，$D_{ij}(t)$表示SFC_i上VNF_j传输到其孪生服务节点的数据量。其次，关联链路传播时延$\tau_{mn,\mathrm{pro}}^{i,j}(t)$则主要与物理链路的距离有关，其计算公式为$\tau_{mn,\mathrm{pro}}^{i,j}(t)=d_{mn}/c$，其中$d_{mn}$表示物理链路$mn$之间的距离。最后，数字孪生构建时延$\tau_{\mathrm{con}}^{i,j}(t)$则与VNF的数据量和孪生服务节点的计算能力有关，其计算公式为：

$$\tau_{\mathrm{con}}^{i,j}(t)=\frac{D_{ij}(t)}{z_{n,l}^{i,j}(t)\cdot c_l^{i,j}(t)\cdot\mu} \tag{7}$$

式中，μ表示孪生服务节点对数据包的处理速率系数。综上所述，数字孪生同步时延为关联链路传输时延、关联链路传播时延与数字孪生构建时延之和。因此SFC_i上VNF_j的数字孪生同步时延$\tau^{i,j}(t)$可表示为：

$$\tau^{i,j}(t)=z_{s,l}^{i,j}(t)\cdot\sum_{mn\in L_{s,l}^{i,j}}[\tau_{mn,\mathrm{tran}}^{i,j}(t)+\tau_{mn,\mathrm{pro}}^{i,j}(t)]+\tau_{\mathrm{con}}^{i,j}(t) \tag{8}$$

1.2 优化目标

为了满足新型电力系统高可靠、低延迟的服务需求，本文通过构建VNF数字孪生体来采集实时的VNF资源需求信息，同时设计动态的VNF映射方案和孪生服务节点关联方案，在最大限度减少服务端到端延迟的同时，降低VNF的数字孪生同步时延。其中数字孪生同步时延越低，表明SDN控制器能够及时获得实时的VNF资源需求信息设计VNF映射方案，防止VNF映射方案由于SDN控制器采集VNF资源需求信息不及时而造成VNF映射策略滞后。

本文考虑的优化问题如下：在保障用户时延最小化的同时降低VNF到其数字孪生体的数字孪生同步时延，并将物理网络服务器的CPU计算资源、存储资源、物理链路的带宽资源及SFC的端到端延迟作为约束条件，设计了一个效用函数$U(t)$，其可表示为：

$$U(t)=-\eta_1\frac{\sum_{i\in F}d^i(t)}{\sum_{i\in F}D^i(t)}-\eta_2\frac{\sum_{i\in F,j\in N_{\mathrm{V}}^i}\tau^{i,j}(t)}{\sum_{i\in F,j\in N_{\mathrm{V}}^i}x_n^{i,j}(t)\cdot\tau_{\max}} \tag{9}$$

式中，η_1和η_2表示两个权重值，为了平衡端到端延迟和VNF平均数字孪生同步时延对全局效用的影响，设两者的权重相同，即$\eta_1=\eta_2=0.5$；$D^i(t)$表示第i条SFC的端到端延迟要求；$\tau_{\max}$表示所有SFC的所有VNF中数字孪生同步时延的最大值，目的是将数字孪生同步时延与端到端延迟做归一化处理，保证两者处在同一量级。

本文的优化目标是通过设计合理的VNF映射方案和孪生服务节点关联方案以最大化该效用函数。由于只有确定了VNF映射位置才能对VNF进

行孪生服务节点关联，将问题分为VNF映射和孪生服务节点关联两个阶段进行求解。其中C1～C7为VNF映射阶段需要满足的约束，C8～C12为孪生服务节点关联阶段的约束条件。优化目标与相关约束可表示为：

$$\max_{x_n^{i,j}(t),\,y_{mn}^{i,jk}(t),\,z_{n,l}^{i,j}(t)} U(t)$$

$$\text{s.t. C1:}\sum_{n\in N_{\mathrm{P}}^{\mathrm{N}}} x_n^{i,j}(t)=1,\forall n\in N_{\mathrm{P}}^{\mathrm{N}}$$

$$\text{C2:}\sum_{i\in F}\sum_{j\in N_{\mathrm{V}}^i} x_n^{i,j}(t)\cdot C_n^{i,j}(t)\leqslant C_n,\forall n\in N_{\mathrm{P}}^{\mathrm{N}}$$

$$\text{C3:}\sum_{i\in F}\sum_{j\in N_{\mathrm{V}}^i} x_n^{i,j}(t)\cdot M_n^{i,j}(t)\leqslant M_n,\forall n\in N_{\mathrm{P}}^{\mathrm{N}}$$

$$\text{C4:}\sum_{i\in F}\sum_{j,k\in N_{\mathrm{V}}^i} y_{mn}^{i,jk}(t)\cdot B_{mn}^{i,jk}(t)\leqslant B_{mn},\forall mn\in L_{\mathrm{P}}$$

$$\text{C5:}\ d^i(t)<D^i(t),\forall i\in F$$

$$\text{C6:}\ x_n^{i,j}(t)=\{0,1\},\forall i\in F,\forall j\in N_{\mathrm{V}}^i,\forall n\in N_{\mathrm{P}}^{\mathrm{N}}$$

$$\text{C7:}\ y_{mn}^{i,jk}(t)=\{0,1\},\forall i\in F,\forall j\in N_{\mathrm{V}}^i,\forall mn\in L_{\mathrm{P}}$$

$$\text{C8:}\sum_{l\in N_{\mathrm{P}}^{\mathrm{DT}}} z_{s,l}^{i,j}(t)=1,\forall s\in N_{\mathrm{P}}^{\mathrm{N}},\forall l\in N_{\mathrm{P}}^{\mathrm{DT}}$$

$$\text{C9:}\sum_{i\in F}\sum_{j\in N_{\mathrm{V}}^i} z_{s,l}^{i,j}(t)\cdot c_l^{i,j}(t)\leqslant C_l^{\mathrm{DT}},\forall l\in N_{\mathrm{P}}^{\mathrm{DT}}$$

$$\text{C10:}\sum_{i\in F}\sum_{j\in N_{\mathrm{V}}^i} z_{s,l}^{i,j}(t)\cdot m_l^{i,j}(t)\leqslant M_l^{\mathrm{DT}},\forall l\in N_{\mathrm{P}}^{\mathrm{DT}}$$

$$\text{C11:}\sum_{i\in F}\sum_{j\in N_{\mathrm{V}}^i} z_{s,l}^{i,j}(t)\cdot F\left[mn\in L_{s,l}^{i,j}\right]\cdot b_{mn}^{i,j}(t)+\sum_{i\in F}\sum_{j,k\in N_{\mathrm{V}}^i} y_{mn}^{i,jk}(t)\cdot B_{mn}^{i,jk}(t)\leqslant B_{mn},\forall mn\in L_{\mathrm{P}}$$

$$\text{C12:}\ z_{s,l}^{i,j}(t)=\{0,1\},\forall i\in F,\forall j\in N_{\mathrm{V}}^i,\forall s\in N_{\mathrm{P}}^{\mathrm{N}},\forall l\in N_{\mathrm{P}}^{\mathrm{DT}} \tag{10}$$

其中，约束C1用于保证虚拟网络中每个VNF只能选择一个物理服务器进行映射。C2和C3确保每台服务器分配给映射在其上VNF的计算资源和存储资源不超过该通用服务器的资源容量。C4确保映射到某条物理链路的带宽资源不超过该物理链路的总带宽资源。C5用于保证每条SFC都满足其时延要求。C6和C7分别表示VNF和虚拟链路映射的二进制变量。C8用于保证虚拟网络中每个VNF只能选择一个孪生服务节点关联。C9和C10确保每台服务器分配给构建在其上的VNF数字孪生体的计算资源和存储资源不超过该服务器的资源容量。C11确保映射到某条物理链路的带宽资源不超过该物理链路的总带宽资源。C12为VNF孪生服务节点关联的二进制变量。

2　基于深度强化学习的VNF映射与孪生服务节点关联算法

2.1　MDP模型

本文将优化问题转化为一个MDP模型，该MDP模型主要包含状态空间、动作空间和回报函数。由于原问题被拆分为VNF映射阶段和孪生服务节点关联阶段，下面将分别介绍这两个阶段的状态空间、动作空间与回报函数。

（1）VNF映射阶段：

在VNF映射阶段，设状态空间为S_1，定义$s_1(t)$为网络在t时隙的状态，其由每条SFC的队列状态、NFV节点的计算资源、存储资源和链路的带宽资源组成，可表示为：

$$s_1(t)=\{R^i(t),C_n(t),M_n(t),B_{mn}(t)\},\ \forall i\in F,\forall n\in N_{\mathrm{P}}^{\mathrm{N}},\forall mn\in L_{\mathrm{P}} \tag{11}$$

设动作空间为A_1，定义$a_1(t)$为网络在t时隙VNF映射阶段采取的动作，主要包括VNF的映射动作$x_n^{i,j}(t)$和虚拟链路的映射动作$y_{mn}^{i,jk}(t)$，可表示为：

$$a_1(t)=\{x_n^{i,j}(t),y_{mn}^{i,jk}(t)\},\forall i\in F,\forall j,k\in N_{\mathrm{V}}^i,\forall mn\in L_{\mathrm{P}} \tag{12}$$

同时设奖励函数为VNF映射优化的端到端延迟，即：

$$r_1(t)=\sum_{i\in F} d^i(t)\Big/\sum_{i\in F} D^i(t) \tag{13}$$

（2）孪生服务节点关联阶段：

在孪生服务节点关联阶段，设状态空间为S_2，定义$s_2(t)$为网络在t时隙完成VNF映射后的状态，其由孪生服务关联节点的计算资源、存储资源和链路的带宽资源组成，可表示为：

$$s_2(t)=\{C_l^{\mathrm{DT}}(t),M_l^{\mathrm{DT}}(t),B_{mn}(t)\},\ \forall l\in N_{\mathrm{P}}^{\mathrm{DT}},\forall mn\in L_{\mathrm{P}} \tag{14}$$

设动作空间为A_2，定义$a_2(t)$为网络在t时隙孪生服务节点关联阶段采取的的动作，主要包括孪生服务节点的关联动作$z_{s,l}^{i,j}(t)$，可表示为：

$$a_1(t)=\{z_{s,l}^{i,j}(t)\},\forall i\in F,\forall j\in N_{\mathrm{V}}^i,\forall s\in N_{\mathrm{P}}^{\mathrm{N}},\forall l\in N_{\mathrm{P}}^{\mathrm{DT}} \tag{15}$$

同时设奖励函数为总的效用函数，即：

$$r_2(t)=U(t) \tag{16}$$

2.2　基于DPPO的VNF映射算法

在VNF映射阶段，采用DPPO算法进行求解。PPO算法利用新旧策略比率限制了动作的更新幅

度，使得算法能够稳健训练收敛到最佳。A3C 算法采用异步学习的方式探索多个环境，在加快模型训练速度的同时提高了算法的收敛性能。DPPO算法结合了 PPO 算法的性能优势与 A3C 算法的异步优势，使得算法的收敛性能更好，训练速度更快。

由于 DPPO 算法采用了 A3C 多智能体的学习框架，与 A3C 算法一样，它包含一个全局网络和多个本地网络，各本地网络与全局网络一致，都采用 PPO 的算法结构。每个本地网络相当于一个智能体，可以独立地探索网络环境进行模型训练，并定期向全局 PPO 网络推送梯度进行网络参数的更新。DPPO 算法的详细过程为：

（1）首先，算法将初始化全局 Actor 和 Critic 网络参数为 (θ_a,θ_c)，全局网络迭代次数为 E，局部网络迭代次数 M、线程数为 N，学习率为 (α_a,α_c)。其次，各个智能体与本地环境交互，从本地环境中获得 SFC 的队列信息及物理节点、链路的资源情况。然后，智能体将采取相应的 VNF 映射动作，环境执行动作后会给予智能体一定的奖励回报，该奖励称为瞬时奖励。最后，各智能体与本地环境不断交互获得奖励来更新 VNF 映射策略，并计算本地梯度和参数更新全局网络。对于智能体 n 来说，本地 PPO 网络采用 A-C 算法的网络架构，即每个网络包含 Actor 和 Critic 网络，Actor 网络用于动作的选择，Critic 网络基于状态值函数和优势函数来评估 Actor 网络所选动作的好坏。其中，智能体 n 的状态值函数是指在当前状态 $s_1^n(t)$ 下执行 VNF 映射动作后累计的奖励期望值，其计算公式为：

$$V\left[s_1^n(t)\right]=E_{\pi(a_1^n(t)|s_1^n(t))}\left\{r_1(t)+\gamma_{v1}V\left[s_1^n(t+1)\right]\right\} \tag{17}$$

（2）智能体 n 在当前状态 $s_1^n(t)$ 下采取动作 $a_1^n(t)$，可以根据映射策略得到一个状态动作值函数，用来表示执行动作后的累计期望值，称为智能体 n 的状态动作值函数，因此智能体 n 的状态动作值函数 $Q\left[s_1^n(t),a_1^n(t)\right]$ 可表示为：

$$Q\left[s_1^n(t),a_1^n(t)\right]=E_{\pi(a_1^n(t)|s_1^n(t))}\left\{r_1(t)+\gamma_{v1}Q\left[s_1^n(t+1),a_1^n(t+1)\right]\right\} \tag{18}$$

此外，DPPO 与 PPO 算法一致，通过引入优势函数来评估选取动作的好坏，其中优势函数 $A\left[s_1^n(t),a_1^n(t)\right]$ 可表示为：

$$A\left[s_1^n(t),a_1^n(t)\right]=Q\left[s_1^n(t),a_1^n(t)\right]-V\left[s_1^n(t)\right] \tag{19}$$

（3）由于基于裁剪优化目标的 PPO 算法优于自适应散度惩罚系数的 PPO 算法[11]，本文采用基于裁剪优化目标来计算 Actor 网络的损失函数，并用于 Actor 网络参数的更新。智能体 n 基于裁剪优化目标的损失函数为：

$$L_{clip}(\theta_a^n)=E_t\left\{\min\left\{c(\theta_a^n)A\left[(s_1^n(t),a_1^n(t)\right],\ clip\left[c(\theta_a^n),1-\rho,1+\rho\right]\cdot A\left[s_1^n(t),a_1^n(t)\right]\right\}\right\} \tag{20}$$

式中，$c(\theta_a^n)=\pi\left[s_1^n(t)\,|\,a_1^n(t),\theta_a^n\right]\Big/\pi_{\text{old}}\left[s_1^n(t)\,|\,a_1^n(t),\theta_a^n\right]$ 表示新旧策略的比值，用来限制策略更新的幅度；$clip(\cdot)$ 表示裁剪函数，用来将新旧策略的比值限制在 $(1-\rho,1+\rho)$ 之间。智能体 n 中 Actor 网络的损失函数可用于全局 Actor 网络梯度和参数的更新，梯度和参数的具体的更新公式为：

$$\Delta\theta_a=\Delta\theta_a+\frac{\partial L_{clip}(\theta_a^n)}{\partial\theta_a^n} \tag{21}$$

$$\theta_a=\theta_a+\alpha_a\Delta\theta_a \tag{22}$$

Critic 网络则采用基于 TD 误差的损失函数进行网络更新，损失函数为：

$$L(\theta_c^n)=\left\{r_1(t)+\gamma_{v1}r_1(t+1)+\cdots+\gamma_{v1}^{T-1}r_1(T+1)-V\left[s_1^n(t)\right]\right\}^2 \tag{23}$$

同理，全局 Critic 网络梯度和参数更新公式为：

$$\Delta\theta_c=\Delta\theta_c+\frac{\partial L(\theta_c^n)}{\partial\theta_c^n} \tag{24}$$

$$\theta_c=\theta_c+\alpha_c\Delta\theta_c \tag{25}$$

基于 DPPO 的 VNF 映射算法框架如图 2 所示。

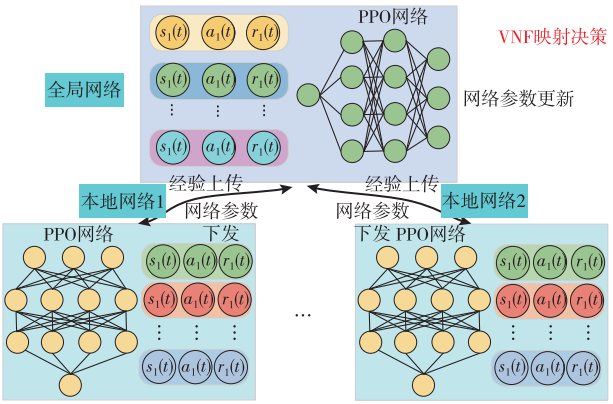


图 2 基于 DPPO 的 VNF 映射算法框架

2.3 基于 A2C 的孪生服务节点关联算法

在孪生服务节点关联阶段，本文采用 A2C 算法进行求解。A2C 算法基于 Actor-Critic 网络架构进行策略选择，但相比于 AC 算法，它又采用了基线的

方法来降低。A2C算法的详细过程如下：

（1）首先，算法将初始化Actor、Critic网络的网络参数和学习率分别为(μ_a,μ_c)和(β_a,β_c)。然后，通过观测VNF映射后的网络环境可以得到孪生服务节点关联阶段的状态$s_2(t)$，并基于Actor网络策略$\pi_2\left[a_2(t)\middle|s_2(t);\mu_a\right]$选择相应的关联动作$a_2(t)$，在智能体执行动作后，智能体将从环境得到一个奖励$r_2(t)$，同时环境也将进入到下一状态再次做VNF映射决策。紧接着，智能体可以根据$t+1$时隙Actor网络的策略$\pi_2\left[\tilde{a}_2(t+1)\middle|s_2(t+1);\mu_a\right]$选择动作$\tilde{a}_2(t+1)$，但不执行该动作，同时让Critic网络打分；最后Actor网络和Critic网络可以根据损失函数计算相应的梯度来更新网络参数。基于A2C的孪生服务节点关联算法如表1所示。

表1　基于A2C的孪生服务节点关联算法

输入：物理网络图G_P和VNF映射策略 输出：SFC上各VNF的孪生服务节点关联策略
1. 初始化Actor和Critic网络参数为(μ_a,μ_c)，同时初始化Actor和Critic网络学习率为(β_a,β_c)
2. 从环境观测当前状态及当前时隙VNF的映射策略，根据Actor网络$\pi_2\left[a_2(t)\middle|s_2(t);\mu_a\right]$选择动作$a_2(t)$，并执行该动作
3. 智能体从环境得到一个奖励值$r_2(t)$，环境由此进入到下一状态
4. 根据Actor网络策略$\pi_2\left[\tilde{a}_2(t+1)\middle|s_2(t+1);\mu_a\right]$选择$t+1$时隙的动作$\tilde{a}_2(t+1)$，但不执行该动作
5. Critic网络对Actor网络t和$t+1$时隙所选的动作进行打分，输出的Q值分别为$\hat{q}_2(t)=q\left[s_2(t),a_2(t);\mu_a\right]$和$\hat{q}_2(t+1)=q\left[s_2(t+1),\tilde{a}_2(t+1);\mu_a\right]$
6. Actor网络通过以下方式更新网络参数：$\mu_a\leftarrow\mu_a-\beta_a\cdot\hat{q}(t)\cdot\nabla_{\mu_a}\ln\pi\left[a_2(t)\middle|s_2(t);\mu_a\right]$
7. Critic网络通过以下方式更新网络参数：$\mu_c\leftarrow\mu_c-\beta_c\cdot\delta(t)\cdot\nabla_{\mu_c}q\left[s_2(t),a_2(t);\mu_c\right]$

（2）Critic网络通过输出Q值$q\left[s_2(t),a_2(t);\mu_a\right]$对智能体选择的动作进行评估。在$t$时隙当智能体执行完动作$a_2(t)$后，Critic网络可以输出$\hat{q}_2(t)=q\left[s_2(t),a_2(t);\mu_a\right]$表示选择动作的好坏，同理在$t+1$时隙当Actor网络选择动作后，Critic网络通过输出$\hat{q}_2(t+1)=q\left[s_2(t+1),\tilde{a}_2(t+1);\mu_a\right]$对其进行评估。

（3）Critic网络可以根据t和$t+1$时隙的Q值计算TD目标值与相应的TD误差，用于后续网络参数的更新。其中TD目标值计算公式为：

$$y_2(t)=r_2(t)+\hat{q}(t+1) \tag{26}$$

TD误差指TD目标值与Critic网络输出Q值之间的差值，其计算公式为：

$$\delta(t)=\hat{q}(t)-y_2(t) \tag{27}$$

在网络参数更新阶段，Actor网络通过以下方式对网络参数进行更新：

$$\mu_a\leftarrow\mu_a-\beta_a\cdot\hat{q}(t)\cdot\nabla_{\mu_a}\ln\pi\left[a_2(t)\middle|s_2(t);\mu_a\right] \tag{28}$$

Critic网络通过以下方式对网络参数进行更新：

$$\mu_c\leftarrow\mu_c-\beta_c\cdot\delta(t)\cdot\nabla_{\mu_c}q\left[s_2(t),a_2(t);\mu_c\right] \tag{29}$$

3　性能仿真与分析

为了评估所提算法的收敛性及有效性，实验将SFC端到端时延、VNF数字孪生同步时延和网络效用等作为评估标准，对算法进行了仿真评估。同时为了表明所提算法的优势，与文献[12]基于遗传的VNF部署算法、文献[13]中DQN算法和文献[11]的PPO算法进行了对比。由于上述算法均未考虑孪生服务节点的关联问题，在对比孪生服务节点关联方案时，本文选取AC算法作为对照。所有仿真实验均在搭载i7 CPU、16GB内存的主机上运行，仿真环境为Python 3.7，并基于Pytorch框架搭建神经网络。

3.1　网络参数设置

设物理网络为全连接网络，包括10台通用服务器，其中6台作为NFV节点，剩下4台作为孪生服务节点，假设每条SFC由2～5个VNF组成，其数据包到达过程服从的泊松分布，数据包大小为500 kByte/packet。并设每一个VNF到其孪生服务节点的数据包大小为100 kByte/packet，并参考文献[14]对网络的相关参数进行设置，其他网络仿真参数如表2所示。

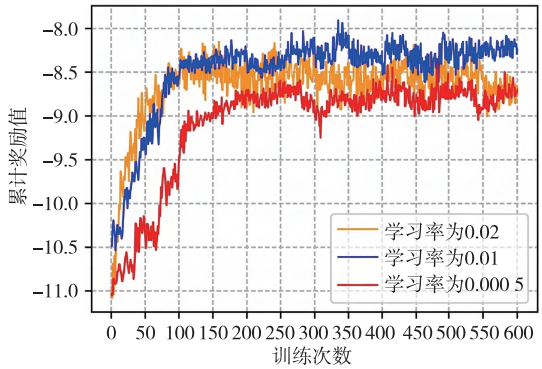
表2　网络仿真参数设置

仿真参数	数值
SFC数据包到达过程	泊松过程$\lambda_i=2$
SFC数据包大小	400～600 kByte/packet
VNF数字孪生数据包大小	50～150 kByte/packet
通用服务器数量	10（6台NFV，4台DT）
SFC时延限制	30 ms
通用服务器CPU资源	8核
通用服务器存储资源	300～100 Mbps
链路带宽资源	640 MB
SFC长度	2～5

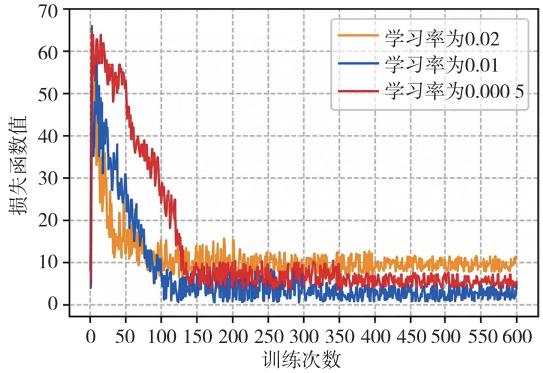
3.2 仿真结果分析

3.2.1 VNF映射阶段仿真结果分析

实验设置数据包到达服从泊松分布，所以数据包的到达是随机的，导致VNF与虚拟链路的映射空间维度较高。因此本文采用DPPO算法求解，通过与不同环境交互得到近似最优的VNF映射策略，并基于最优的映射策略采用A2C算法求得近似最优的孪生服务节点关联策略。图3分别展示了VNF映射阶段所提算法在不同Actor网络学习率下的累计奖励和损失函数，证明了所提基于DPPO的VNF映射算法的收敛性，其中Critic网络学习率固定取值为0.001。



（a）DPPO算法累计奖励值对比



（b）DPPO算法损失函数值对比

图3 DPPO算法累计奖励值与损失函数值对比

从图中可以看出，DPPO算法在不同Actor网络学习率下的累计奖励有所差异，当学习率为0.002时算法收敛最快，学习率为0.000 5时算法收敛最慢。这是因为学习率较高时算法的动作策略幅度较大能加快算法收敛，学习率较低时算法的动作策略幅度较小导致算法收敛较慢。而当学习率为0.001时算法获得的累积奖励较高，因此本文设置DPPO算法的学习率为0.001。

图4展示了在不同SFC数量下，所提算法与其他基准算法在系统总时延方面的性能对比。

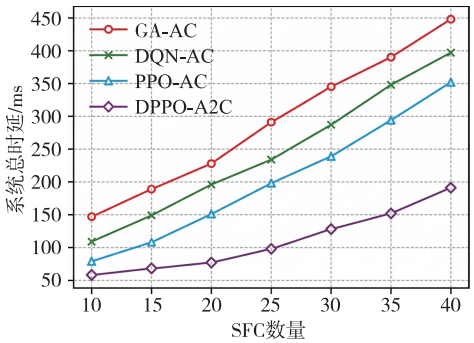
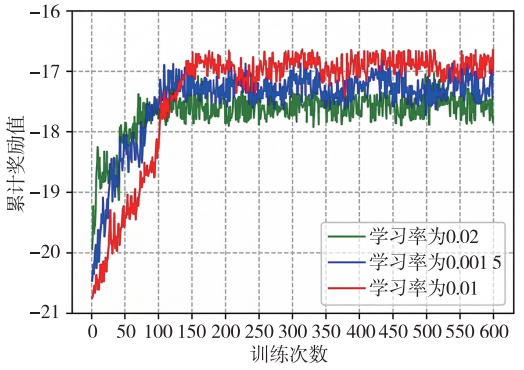


图4 系统总时延对比

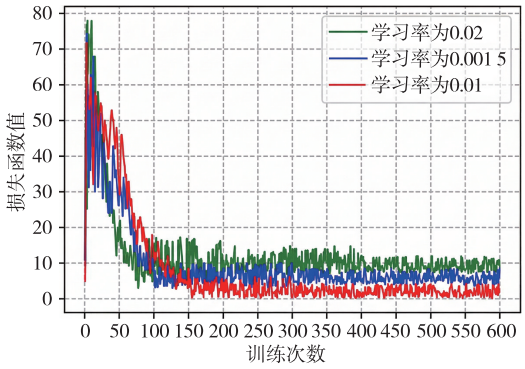
从图中可以看出，随着SFC数量的增加，所有算法的系统总时延都将增大。其次，GA-AC算法在系统总时延方面效果较差，这是因为其中的GA算法是一种启发式算法，随着SFC数量的增加使得VNF映射需要搜索的解空间变得更大，因此GA算法在SFC数量较大的环境下极易陷入局部最优，导致系统总时延最大。DQN-AC算法中的DQN算法作为基本的离散动作强化学习算法，相比于GA算法系统总时延得到了一定的下降，但VNF映射不稳定导致无法合理有效地利用网络资源。PPO-AC算法中的PPO算法通过新旧策略比率来选择VNF映射动作，极大限制了动作更新的幅度，使得VNF映射得更加稳定，所以系统时延得到了有效降低。而DPPO-A2C算法中的DPPO算法通过多个智能体对环境进行了全面搜索，能寻找到更为丰富多样的VNF映射解空间，因此随着SFC数量和VNF映射动作增多，系统能获得最低的总时延，且与其他算法相比优势逐渐增大。

3.2.2 孪生服务节点关联阶段仿真结果分析

设孪生服务节点关联阶段DPPO-A2C算法中A2C算法的Critic网络学习率为0.001，通过调整Actor网络学习率来获得全局最优解。图5展示了A2C中算法在不同Actor网络学习率下的累计奖励与损失函数值对比。由图可知，当A2C算法的Actor网络学习率为0.002时，算法收敛得最快，但抖动也最明显，导致网络的损失函数值陷入局部最优。而随着学习率的下降，算法的收敛速率变慢，但收敛得更加稳定。当Actor网络学习率为0.001时，算法获得得奖励最大，因此实验设A3C算法Actor网络的学习率为0.001。



（a）A2C 算法累计奖励值对比



（b）A2C 算法损失函数值对比

图 5　A2C 算法累计奖励值与事损失函数值对比

图 6 展示了在不同 SFC 数量下，所提算法与其他基准算法在 VNF 平均数字孪生同步时延方面的性能对比。

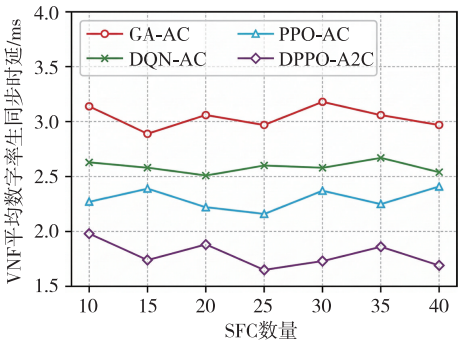


图 6　VNF 平均数字孪生同步时延对比

从图中可以看出，GA-AC 算法的 VNF 平均数字孪生同步时延最大，这是因为 GA 算法作为启发式算法，导致 VNF 映射方案不合理，陷入局部最优，因此关联的同步时延最大；DQN-AC 和 PPO-AC 算法中的 DQN 和 PPO 算法 VNF 映射更好，同步时延有所降低。DPPO-A2C 中的 DPPO 算法通过多智能体对全局 VNF 映射方案进行了全面搜索，且相比于 AC 算法，A2C 算法通过优势函数代替 Critic 网络中的回报函数，作为衡量选取动作值相较于所有动作平均值的好坏，因此获得了最低的 VNF 平均数字孪生同步时延。

通过 SFC 时延值与数字孪生同步时延值来衡量系统的效用，图 7 展示了在不同 SFC 数量下，本文算法与其他基准算法的效用对比。

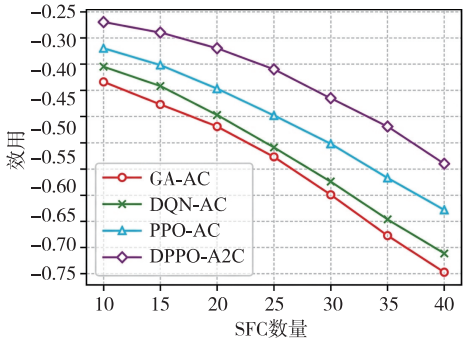


图 7　效用对比

由图可知，随着 SFC 数量的增加，所有 SFC 的总时延和 VNF 的数字孪生同步时延都将增大，因此系统总时延会减小。其中 GA-AC 算法的效用值下降得最快，而所提 DPPO-A2C 算法通过 DPPO 进行全局搜索，同时又借助优势函数评估 A2C 所选动作的优劣，所以效用值最高且下降得最慢。

4　结束语

针对新型电力系统中服务的动态变化特性和低延迟需求，提出了一种数字孪生辅助的网络切片资源分配方案。首先，通过构建 VNF 数字孪生体来捕捉动态变化的 VNF 资源需求，并将数字孪生辅助的 VNF 映射优化问题建立为一个 MDP 模型，该模型以最小化 SFC 部署时延和 VNF 平均数字孪生同步时延为目标，同时受限于 SFC 端到端时延、通用服务器的 CPU 资源、存储资源及物理链路带宽资源等约束。其次，考虑到 VNF 映射与孪生服务节点关联问题的先后性，本文将原问题分为 VNF 映射和孪生服务节点关联两个阶段来求解，并提出了基于 DPPO 的 VNF 映射算法和基于 A2C 的孪生服务节点关联算法。仿真结果表明，所提算法能够通过设计合理的 VNF 映射与关联方案来降低 SFC 的端到端时延和 VNF 的数字孪生同步时延。

参考文献：

[1]　臧玉华，郑焕坤，尹世豪．面向新型电力系统的 5G 网络切片资源分配策略[J]．河北电力技术，2023，42(1)：26-31.
Zang Yuhua，Zheng Huankun，Yin Shihao．5G network slice

resource allocation strategy for new power systems[J]. Hebei Electric Power, 2023, 42(1): 26-31.

[2] Sun P, Lan J, Li J, et al. Combining deep reinforcement learning with graph neural networks for optimal VNF placement [J]. IEEE Communications Letters, 2021, 25(1): 176-180.

[3] 刘夏宇, 王振雄, 易 皓, 等. 数字化主动配电网中光伏逆变器并网主动支撑方案研究[J]. 河北电力技术, 2021, 40(5): 1-5, 19.
Liu Xiayu, Wang Zhenxiong, Yi Hao, et al. Research on active support scheme of grid-connected photovoltaic inverters in digital active distribution network[J]. Hebei Electric Power, 2021, 40(5): 1-5, 19.

[4] 周志恒, 张 平, 张琳娟, 等. 面向配电网业务的5G切片部署策略研究[J]. 电子设计工程, 2023, 31(4): 28-32, 42.
Zhou Zhiheng, Zhang Ping, Zhang Linjuan, et al. Research on deployment strategy of 5G slice for distribution network service [J]. Electronic Design Engineering, 2023, 31(4): 28-32, 42.

[5] 陈 俊, 黄飞宇, 黎作明. 基于DQN的电力物联网5G边缘切片资源管理研究[J]. 电测与仪表, 2022, 59(1): 155-161.
Chen Jun, Huang Feiyu, Li Zuoming. Research on DQN-based 5G edge slicing resource management of power Internet of Things [J]. Electrical Measurement & Instrumentation, 2022, 59(1): 155-161.

[6] 王井龙. 网络切片下基于遗传算法的虚拟网资源分配算法[J]. 江苏通信, 2021, 37(6): 22-25.

[7] 蒲天骄, 陈 盛, 赵 琦, 等. 能源互联网数字孪生系统框架设计及应用展望[J]. 中国电机工程学报, 2021, 41(6): 2012-2028.
Pu Tianjiao, Chen Sheng, Zhao Qi, et al. Framework design and application prospect for digital twins system of energy Internet[J]. Proceedings of the CSEE, 2021, 41(6): 2012-2028.

[8] Mittal S, Ahmad Khan M, Romero D, et al. Smart manufacturing: characteristics, technologies and enabling factors [J]. Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture, 2019, 233(5): 1342-1361.

[9] 中国移动研究院. 数字孪生网络白皮书[R/OL]. (2021-09) [2026-02-25]. https://max.book118.com/html/2021/1002/7113025043004014.shtm
China Mobile Research Institute Digital twin network white paper [R/OL]. (2021-09) [2026-02-25]. https://max.book118.com/html/2021/1002/7113025043004014.shtm

[10] 唐 伦, 杨 恒, 马润琳, 等. 基于5G接入网络的多优先级虚拟网络功能迁移开销与网络能耗联合优化算法[J]. 电子与信息学报, 2019, 41(9): 2079-2086.
Tang Lun, Yang Heng, Ma Runlin, et al. Multi-priority based joint optimization algorithm of virtual network function migration cost and network energy consumption[J]. Journal of Electronics & Information Technology, 2019, 41(9): 2079-2086.

[11] Schulman J, Wolski F, Dhariwal P, et al. Proximal policy optimization algorithms [J]. arXiv, 2017: arXiv: 1707.06347.

[12] 陈 卓, 冯 钢, 刘 蓓, 等. 运营商网络中面向时延优化的服务功能链迁移重配置策略[J]. 电子学报, 2018, 46(9): 2229-2237.
Chen Zhuo, Feng Gang, Liu Bei, et al. Delay optimization oriented service function chain migration and re-deployment in operator network[J]. Acta Electronica Sinica, 2018, 46(9): 2229-2237.

[13] Xu S, Li Y, Guo S, et al. Cloud - edge collaborative SFC mapping for industrial IoT using deep reinforcement learning[J]. IEEE Transactions on Industrial Informatics, 2022, 18(6): 4158-4168.

[14] 唐 伦, 杨 恒, 赵国繁, 等. 基于时延感知的5G网络切片节点和链路映射算法[J]. 北京邮电大学学报, 2018, 41(6): 71-77.
Tang Lun, Yang Heng, Zhao Guofan, et al. Delay-aware 5G network slicing node and link embedding algorithm[J]. Journal of Beijing University of Posts and Telecommunications, 2018, 41(6): 71-77.

作者简介：

唐奕明(2001—),男,硕士研究生,研究方向为电力系统暂态稳定性分析。